APPARATUS AND METHOD FOR ADAPTIVE SPATIAL SEGMENTATION-BASED NOISE REDUCING FOR ENCODED IMAGE SIGNAL

BACKGROUND OF THE INVENTION

Field of the Invention:

[0001] The invention relates to image noise reduction techniques primarily operable in real-time by apparatus and methods for reducing the correlated noise in an image or a sequence of images. More particularly, the invention relates mainly to spatial adaptive techniques for mosquito noise reduction in Discrete Cosine Transform (DCT) based decoded image applications.

Description of the Prior Art:

[0002] Recently, many international standards for still image and video compression such as the ITU-T H261, H263, and the ISO JPEG, MPEG-1, MPEG-2 standards have mainly proposed the block based Discrete Cosine Transform (DCT) as a possible compression technique.

[0003] At low and moderate bit rates, block-based DCT coding artifacts become perceptible. Such artifacts are known as mosquito noise or ringing noise occurring around edges within an image or near a smooth zone as well as the blocking effect. For still pictures or still parts of image, the blocking effect is dominant and visible in smooth regions. For dynamic video sequences, mosquito noise becomes more evident for the human vision system (HVS) than the blocking effect.

[0004] There are many existing techniques for blocking effect reduction. In H. Reeve and J. Lim, “Reduction of blocking effects in image coding”, Optical Engineering, vol. 23, Jan/Feb 1984, pp. 34-37, the authors teach the systematical use of low-pass filters applied at block boundary. Low pass filtering is utilized also in U.S. Patent No. 5,850,294 to Apostolopoulos et al. for blocking artifact reduction purposes. However, the blocks that potentially exhibit block artifacts are detected in the DCT domain and low-pass filtering is applied only for the distorted blocks. In B. Ramamurthi and A. Gersho, “Nonlinear Space-variant post processing of block coded images”, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-34, Oct 1986, pp. 1258-1268, the proposed adaptive filtering is based on the detection of edge orientation at each block boundary pixel. Many authors, as in, for instance, A. Zakhor,

"Iterative Procedure for Reduction of Blocking Effects in Transform Image Coding", IEEE Transactions on Circuits and Systems for Video Technology, vol. 2, No.1, Mar 1992, pp. 91-95, have proposed various multi-pass procedure techniques for this purpose. The iterative techniques can provide potentially a higher performance than the non-iterative ones, but are less attractive for real time processing.

[0005] For mosquito noise artifact reduction (MNR), in U.S. Patent No. 5,610,729, Nakajima teaches an estimation of block mean noise using the quantization step and the I, P, B coding mode when these data are available from the compressed bit stream. Nakajima teaches also the use of the well-known Minimum Mean Square Error (MMSE) filter proposed originally by J. S. Lee in "Digital image enhancement and noise filtering by use of local statistics", IEEE Transactions on PAMI-2, Mar 1980, pp. 165-168, for artifact reduction. However, in many applications, the quantization step or the coding mode is not necessary known or accessible. Moreover, while the MMSE filter is efficient for edge reservation, it is not necessary for noise reduction near an edge.

[0006] In U.S. Patent No. 5,754,699, Sugahara proposes a similar approach by using block quantization step size information for noise power estimation and an empiric coring technique for artifact filtering.

[0007] Also for MNR, in U.S. Patent No. 5,850,294, Apostolopoulos et al. propose a filtering on the true non-edge pixels within blocks containing edge pixels rather than smoothing the edge pixels, to avoid eventual blur and picture sharpness loss due to true edge filtering. However, the filtering technique for non-edge pixels is not clearly specified.

[0008] In a same manner, in U.S. Patent No. 5,852,475, Gupta et al. apply separable low pass filters only on portions of an image that are not part of an edge and are not part of areas of texture or fine detail. The proposed post processor contains also a look up table based temporal digital noise reduction unit for reliable edge detection. For the chrominance signals Gupta et al. teach the use of simple low pass filtering. U.S. Patent No. 5,920,356 to Smita et al. is an ameliorated version of U.S. Patent No. 5,852,475 in which the filtering is controlled by a coding parameter of the replenished macro-blocks.

[0009] In U.S. Patent No. 6,064,776 to Kikuchi et al., in a similar manner, a given block is classified according to whether it is considered part of a flat domain or not. If a

block is considered as part of a flat domain, block pixel correction is then given by an AC component prediction technique.

[0010] In U.S. Patent No. 6,188,799, Tan et al. teach the use of separable low-pass filtering, when block boundaries are located, for a serial reduction of blocking effect and then, mosquito noise. For detected blocking effect, the pixels are firstly corrected by a proposed modified version of bilinear interpolation and secondly, by a mean value of homogenous neighboring pixels within the quantization step size.

SUMMARY OF THE INVENTION:

[0011] The present invention provides an apparatus and method for efficiently reducing noise in a block-based decoded image signal.

[0012] According to an aspect of the present invention, there is provided an apparatus for reducing noise in a block-based decoded image signal including a luminance component. The apparatus comprises an image region classifier responsive to said luminance component for analyzing each luminance pixel value of the luminance component according to a corresponding luminance pixel spatial context in a same frame of said image signal to classify the luminance pixel in a selected one of a plurality of predetermined image region classes associated with distinct image region spatial characteristics and to generate a corresponding selected region class indicative signal. The apparatus further comprises a shape-adaptive luminance noise power estimator responsive to said luminance component and said selected region class indicative signal for estimating statistical characteristics of said luminance pixel by using local window segmentation data associated with the luminance pixel, to generate a corresponding luminance noise power statistical characteristics indicative signal; and a shape-adaptive luminance noise reducer for filtering said luminance component according to said luminance noise power statistical characteristics indicative signal. Conveniently, the distinct image region spatial characteristics include edge, near edge, flat, near flat and texture spatial characteristics. Preferably, the block-based decoded image signal further includes first and second chrominance components, and the apparatus further comprises a shape-adaptive chrominance noise power estimator responsive to said chrominance components and said selected region class indicative signal for estimating statistical characteristics of first and second chrominance pixels associated with said luminance

pixel by using local window segmentation data associated with each said chrominance pixel to generate a corresponding chrominance noise power statistical characteristics indicative signal; and a shape-adaptive chrominance noise reducer for filtering each said chrominance component according to said corresponding chrominance noise power statistical characteristics indicative signal.

[0013] According to a further aspect of the present invention, there is provided a method for reducing noise in a block-based decoded image signal including a luminance component. The method comprises the steps of: i) analyzing each luminance pixel value of said luminance component according to a corresponding luminance pixel spatial context in a same frame of said image signal to classify the luminance pixel in a selected one of a plurality of predetermined image region classes associated with distinct image region spatial characteristics and to generate a corresponding selected region class indicative signal; ii) estimating, from said luminance component and said selected region class indicative signal, statistical characteristics of said luminance pixel by using shape-adaptive local window segmentation data associated with the luminance pixel, to generate a corresponding luminance noise power statistical characteristics indicative signal; and iii) filtering said luminance component according to said luminance noise power statistical characteristics indicative signal. Conveniently, the distinct image region spatial characteristics include edge, near edge, flat, near flat and texture spatial characteristics. Preferably, the block-based decoded image signal further includes first and second chrominance components and, method further comprises the steps of: iv) estimating, from said chrominance components and said selected region class indicative signal, statistical characteristics of first and second chrominance pixels associated with said luminance pixel by using shape-adaptive local window segmentation data associated with each said chrominance pixel to generate a corresponding chrominance noise power statistical characteristics indicative signal; and v) filtering each said chrominance components according to said corresponding chrominance noise power statistical characteristics indicative signal.

[0014] According to a further aspect of the present invention, there is provided an apparatus and method for post-processing a decompressed image signal to reduce spatial mosquito noise therein. In particular, the post processor calls for an image multiple region segmentation, region noise power estimations for respectively

luminance and chrominance signal components, and their associated adaptive noise corrections.

[0015] In segmenting an image into regions, the inventive apparatus and method employ edge/no-edge detectors and simple binary consolidation operators to classify and reinforce detected Edge (E), Near-Edge regions (NE), Flat regions (F), Near-flat regions (NF) and finally Texture (T) regions. The preferred segmentation is based essentially on the following observations: First, almost strong mosquito noise is found not only in NE regions but also in NF regions; second, some important noise is also noticeable in picture edges; third, texture masks mosquito noise; and fourth, any excessive filtering in texture or flat regions will degrade eventually fine signal details.

[0016] In estimating local noise power of the luminance component of the image signal, the inventive apparatus and method consider the diagonal high frequency component of the decoded image. The local noise power estimator comprises a local variance calculator that considers only local similar pixels to the current one, a look up table (LUT) for a conversion from observed diagonal high frequency component power to equivalent additive noise power. The noise power estimator also comprises a noise power weighting for each classified region and finally a low-pass filter for smoothing the variation of estimated local noise power between regions. Thus, the proposed method permits different smoothing degree for each segmented region and region transition to ensure resulting image quality.

[0017] For noise correcting, the proposed apparatus and method are based on a shape adaptive local segmented window that considers only the similar intensity pixels to the current one for the local mean and local standard deviation estimations. For reliable window segmentation, a diamond shape two-dimensional (2D) low pass filter is preferably required for the local adaptive windowing. The noise corrector further comprises a gain calculator in order to minimize the Mean Square Error (MMSE) for given local signal mean, local signal power and local additive noise power. The combination of local shape adaptive windowing and MMSE constitutes a noise corrector working on all of the above-cited classified regions.

[0018] It is worthwhile to mention that the proposed mosquito noise filtering also partly reduces the blocking effect.

[0019] From another broad aspect of the present invention, there is also provided an adaptive apparatus and method for noise power estimation and noise correction for

the chrominance components which are severely damaged at low bit rate in a decoded video signal. In estimating local noise power in each chrominance component, the proposed method is similar to luminance component processing. However, in the chrominance case, the region classification is not required. In other words, there is only a single region for the whole image. For noise correcting of the chrominance component, the above luminance-based shape adaptive windowing and the MMSE technique are both utilized in a similar manner to the luminance case. Of course, considering the chrominance-sampling rate requires the use of suitable interpolation and decimation techniques for the chrominance signals.

BRIEF DESCRIPTION OF THE DRAWINGS:

[0020] Embodiments of the present invention will be now described with reference to the accompanying drawings, in which:

[0021] Figure 1 is a general block diagram of a preferred embodiment of a mosquito noise reducing apparatus in accordance with the invention;

[0022] Figure 2 is a block diagram of a Region Classifier (RC) included in the embodiment of Figure 1;

[0023] Figure 3 is a block diagram of a LUminance component Region-Based Noise power Estimator (LU-REBNE) included in the embodiment of Figure 1;

[0024] Figure 4 is a block diagram of a LUminance component LOcal SEGmentation-based Adaptive Noise Reducer (LU-LOSEGANR) included in the embodiment of Figure 1;

[0025] Figure 5 is a block diagram of a CHrominance component LOcal Noise power Estimator (CH-LONE) included in the embodiment of Figure 1;

[0026] Figure 6 is a block diagram of a CHrominance component LOcal SEGmentation-based Adaptive Noise Reducer (CH-LOSEGANR) included in the embodiment of Figure 1;

[0027] Figure 7 is a block diagram of a proposed configuration used for performing an off-line noise variance pre-estimation; and

[0028] Figure 8 illustrates an empirical form of a Look-Up Table (LUT) for a conversion of observed diagonal high frequency component power to equivalent additive noise power.

DETAILED DESCRIPTION:

[0029] Referring now to the drawings, Figure 1 represents a block diagram of an embodiment of a mosquito noise reduction apparatus 50 in accordance with the invention.

[0030] MNR apparatus 50 receives four (4) main system inputs. Image signal luminance Y and chrominance Cu/Cv components are applied at inputs 100 and 101u/v respectively. Coding Parameters at input 102 might represent, for example, an average coding bit rate. In the preferred implementation, this input is simply controlled by an end-user in a heuristic manner. Also, the end-user controlled Signal Mode signal 103 represents the thresholding values pre-determined for an image signal type such as DVD, DSS, DV-25 signal etc.

[0031] MNR apparatus 50 comprises five (5) main blocks: image Region Classifier (RC) 104, LUminance component REgion-Based Noise Power Estimator (LU-REBNE) 106, LUminance component LOcal SEgmentation-based Adaptive Noise Reducer (LU-LOSEGANR) 108, CHrominance component Local Noise power Estimator (CH-LONE) 112 and CHrominance component LOcal SEGgmentation-based Adaptive Noise Reduction (CH-LOSEGANR) 115. It is important to note that, for simplicity, Figure 1 illustrates only one CH-LONE 112 and its associative CH-LOSEGANR 115 for both chrominance components Cu and Cv. Persons of ordinary skill in the art will understand that such components may be implemented in a time sharing manner or in parallel as is well known in the art.

[0032] Image Region Classifier (RC) 104 described in detail below with reference to Figure 2 receives three (3) signals, namely: the decoded luminance Y 100, an interpolated chrominance Cu/Cv 116u/v and the signal mode 103 to generate a region map 105. Image Region Classifier 104 is responsive to the luminance component Y 100 of the decoded image signal for analyzing each luminance pixel value thereof in accordance with a corresponding luminance pixel spatial context in a same frame of said image signal. RC 104 classifies the luminance pixel in a selected one of a plurality of predetermined image region classes associated with distinct image region spatial characteristics and generates a corresponding selected region class indicative signal (Region Map 105). Conveniently, the predetermined image region classes or region map allows the classification of a current pixel as belonging to an edge (E), a

flat region (F), a near flat region (NF), a near edge region (NE) or a finally textured (T) region, as distinct image region spatial characteristics.

[0033] Region map signal 105, luminance signal Y 100 and Coding Parameters 102 are applied as main inputs to the Luminance component REgion-Based Noise power Estimator (LU-REBNE) 106. Two (2) secondary signals 110 and 111 that represent data on the segmented local window generated by the LU-LOSEGANR 108 are also applied to LU-REBNE 106. LU-REBNE is a shape-adaptive luminance noise power estimator that is responsive to the luminance component Y 100 and the selected region class indicative signal (Region Map 105) for estimating statistical characteristics of the luminance pixel by using local window segmentation data associated with the luminance pixel, to generate a corresponding luminance noise power statistical characteristics indicative signal.

[0034] LU-REBNE 106 described further below with reference to Figure 3 yields an estimated luminance noise local standard deviation signal 107 in the decoded luminance component. The noise local standard deviation is required further for a MMSE noise reduction.

[0035] Noise local standard deviation signal 107 and noisy luminance component Y 100 input to LU-LOSEGANR 108 which yields, in turn, a filtered Y luminance signal 109 and the two signals 110 and 111 containing data on the segmented local window characteristics. LU-LOSEGANR is a shape-adaptive luminance noise reducer for filtering the luminance component Y 100 according to the luminance noise power statistical characteristics indicative signal (noise local standard deviation signal 107). LU-LOSEGANR 108 is described further below with reference to Figure 4.

[0036] Chrominance Cu/Cv signals 101u/v, Coding Parameters signal 102 and the segmented local window data signals 110 and 111 are applied to CHrominance component LOcal Noise power Estimator (CH-LONE) 112. CH-LONE 112 provides an estimated chrominance noise local standard deviation signal 113 in the chrominance component, required for a chrominance MMSE noise reduction as is described further below with reference to Figure 5.

[0037] Finally, chrominance noise local standard deviation signal 113 and noisy chrominance Cu/Cv signals 101u/v are input to CH-LOSEGANR 115. CH-LOSEGANR 115 yields, in turn, interpolated chrominance components signals 116u/v optionally

required in the RC block 104, and filtered Cu/Cv chrominance signals 114u/v. CH-LOSEGANR 115 is described further below with reference to Figure 6.

[0038] As is understood by persons of ordinary skill in the art, appropriate delays for signal synchronization required by the various operations of MNR apparatus 50 are not illustrated. Implementation of such delays is well known in the art.

[0039] Referring now to Figure 2, there is illustrated in block diagram Region Classifier (RC) 104 in accordance with the invention.

[0040] A decoded noisy luminance signal Y 100 is applied to the region classifier RC generally designated by 104. Firstly, for a reliable classification, the noisy signal Y is filtered by a diamond shape 2D diagonal low pass filter L1 (201) in order to reduce high frequency noise component. The filter impulse response is given by the following equation:

$$L_1(i,j) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} / 8 \qquad (1)$$

in which the couple (i, j) represents the current coordinates (line, column) of the central and considered pixel. The filter output 202 is sent to four (4) Sobel edge masks 203, 204, 205 and 206 designated respectively for 0°, 90°, 45° and 135°. Their respective impulse responses are:

$$Sobel_0(i,j) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \qquad (2a)$$

$$Sobel_{90}(i,j) = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \qquad (2b)$$

$$Sobel_{45}(i,j) = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} \qquad (2c)$$

$$\text{Sobel}_{135}(i,j) = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}. \tag{2d}$$

[0041] Each of the Sobel masks 203, 204, 205, and 206 has a respective output, 207, 208, 209 and 210, to a respective absolute value detector 211, 213, 215 and 217. The respective outputs 212, 214, 216 and 218 of the detectors 211, 213, 215 and 217 are now utilized for two different purposes: strong edge detection and flat region detection.

[0042] For edge detection, the four (4) absolute value detector outputs 212, 214, 216 and 218 are applied respectively to their associated thresholding (comparison) operators 220, 222, 224 and 226. The thresholding output is equal to 1 if its corresponding input is greater than or equal to a threshold value; otherwise, it will be 0. The pre-determined threshold values at 272 are given by a Look-Up Table (LUT) 273 that is controlled in turn by Signal Mode signal 103. The four comparison operator outputs 229, 230, 231 and 232 are applied together to an OR gate 237 whose output 239 represents a preliminary detection for strong edges in a given image. This detection is far from perfect; the detected edge can be broken or composed of isolated points. To partly remedy the situation, the preliminary detection binary signal 239 is submitted to two (2) non-linear operations in cascade. The first one, Add Only Consolidation (AOC) 241, is defined as follows: Consider a local window centered on the current pixel. If a count of the "1" number in the window is greater than or equal to a threshold, then the operator output is "1"; otherwise, the output remains unchanged. In the preferred implementation, the window dimension is 3x3 and the threshold value, at 240, is set to be 4. The AOC operator can be described by the following:

[0043] Let in(i, j) and out(i, j) denote respectively the input and the output of the operator at the coordinates (i, j) of the current pixel. Let W is the local window domain. The operator output is given by:

$$\text{out}(i, j) = \begin{cases} 1, & \text{if } \sum_{(n,m)\in W} \text{in}(i-n, j-m) \geq \text{Threshold} \\ \text{in}(i, j), & \text{otherwise.} \end{cases} \qquad (3)$$

The second operator 248, Remove Only Consolidation ROC, is in turn given by:

$$\text{out}(i, j) = \begin{cases} 0, & \text{if } \sum_{(n,m)\in W} \text{in}(i-n, j-m) \leq \text{Threshold} \\ \text{in}(i, j), & \text{otherwise.} \end{cases} \qquad (4)$$

In the above equation, in(i, j) and out(i, j) are respectively again the input and the output of the considered operator and W is the local window domain. In other words, if the count of "1" numbers in the window is smaller than or equal to a threshold, then the operator output is "0"; otherwise, the output remains unchanged. In the preferred embodiment, the window dimension is 3x3 and the ROC threshold at 245 is equal to 2. The ROC output signal 251 represents now the detected edge map.

[0044] In order to determine a Near Edge (NE) region, the detected edge map signal 251 is block-based expanded by a binary operator Block-based Add Only Consolidation (BAOC) 253. In the preferred embodiment, the block dimensions are 4 lines by 8 columns. There are a few reasons for these chosen dimensions: first, in some CODECs for recording mediums such as DV-25, DV-50, the block dimension can be 4x8 and in the popular MPEG-2, the compression blocks can be frame-based 8x8 (i.e. in a given field, the dimension of a block is 4x8); second, 4x8, which is a sub-block of 8x8, has been experimentally proved to be a compromise between over-correction and picture naturalness preservation. BAOC 253 is described as follows. In a given block, if the number of edge pixels, represented by a number of "1", is greater than or equal to a threshold, (e.g. 3) at 259, then all pixels in the block become "1"; otherwise, the block remains unchanged. Let B be the considered block domain. The descriptive equation is given by the following Equation (5):

$$\forall (i, j) \in B, \text{out}(i, j) = \begin{cases} 1, & \text{if } \sum_{(i,j)\in B} \text{in}(i, j) \geq \text{Threshold} \\ \text{in}(i, j), & \text{otherwise.} \end{cases} \qquad (5)$$

[0045] BAOC output 258 is then applied to an AND gate 260 together with the negation of binary edge signal 251. Black dots at AND gate inputs denote negation of the considered input in Figure 2. AND gate output 268 from AND gate 260 represents the detected NE region map signal.

[0046] For a flat region detection, the four (4) absolute value detector outputs 212, 214, 216 and 218 are applied respectively to four (4) other associated thresholding (comparison) operators 221, 223, 225 and 227. The thresholding output is equal to 1 if its corresponding input is smaller than a threshold value; otherwise, it will be 0. The pre-determined threshold values at 228 are given also by LUT 273 that is controlled in turn by Signal Mode signal 103. The four comparison operator outputs 233, 234, 235 and 236 are applied to an AND gate 238 whose output 242 represents a preliminary detection for flat regions in a given image. This flat region detection can be composed again of isolated points or isolated holes. To partly remedy the situation, preliminary detection binary signal 242 is submitted to two Add and Remove Conditional Consolidation (ARCC) operators 250 and 259 in series. A complete ARCC operator is given by the following equation:

$$\text{out}(i,j) = \begin{cases} 1, \quad \text{if} \quad \sum\limits_{(m,n)\in W} \text{in}(i-m, j-n) \geq \text{Threshold1} \\ \quad \text{and } \forall(m,n) \in W, | \text{YF}(i-m, j-n) - \text{YF}(i,j) | < \text{Threshold2} \\ \quad \text{and } \forall(m,n) \in W, | \text{CuF}(i-m, j-n) - \text{CuF}(i,j) | < \text{Threshold2} \\ \quad \text{and } \forall(m,n) \in W, | \text{CvF}(i-m, j-n) - \text{CvF}(i,j) | < \text{Threshold2} \\ 0, \quad \text{otherwise.} \end{cases} \qquad (6)$$

[0047] In this equation, signal YF 202 denotes the filtered version of the noisy luminance input 100. Similarly, CuF and CvF at 247u/v correspond to the filtered

version of the interpolated chrominance component inputs Cu and Cv 116u/v as provided by CH-LOSEGANR 115. The filtering is provided by the 2D low pass filters 243u/v. Moreover, in the preferred embodiment, for the first ARCC operator 250, the window dimension is 5x5, threshold1 at 246 is set to 16 and the internal threshold2 is set to 8. For the second ARCC operator 259, the window dimension is 21x21 as empirically chosen for typical video ITU-601 signal, the threshold1 at 255 is set to 3 and the internal threshold2 to 8. The second operator output signal 257 represents the Flat (F) region map.

[0048] It is interesting to note that omitting the chrominance components in Equation (6) yields a possible simplified, but less efficient version for Flat region consolidation.

[0049] The Flat region map signal 257 is applied together with the negation of the first ARCC output 256 to an AND gate 262. The AND gate output 264 represents the corresponding Near-Flat (NF) regions in which mosquito noise is very noticeable for the human vision system (HVS).

[0050] The Texture (T) region in the present embodiment is computed as NOT all of the four (4)-detected regions: E, NE, F and NF. The Texture region map signal 271 can be obtained with a NOT-AND gate 269 with four appropriate corresponding signal inputs: 268, 251, 257 and 262.

[0051] Finally, combining together the five above region maps by the classification block 252 yields the picture Region Map signal 105 utilized for noise power weighting. In order to avoid the potential conflict when a given pixel is classified to more than one region, classification is based on the following priority: Edge, Near Edge, Near Flat, Flat and Texture.

[0052] Referring now to Figure 3, there is illustrated a block diagram for the LUminance component REgion-Based Noise power Estimation (LU-REBNE) generally designated at 106.

[0053] First of all, it can be frequently observed that there is no important signal component in a diagonal high frequency spatial domain. It is thus reasonable to use a diamond shape filter for noise power estimation. Let the noisy decoded luminance signal Y 100 be applied to the diamond shape high pass filter that is composed of a low pass filter 301 whose output 302 is connected to a subtractor 303. Subtractor 303

subtracts output 302 from luminance Y 100. The low pass filter 301 is given by the following impulse response:

$$d3(i,j) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 8 & 2 & 0 \\ 1 & 8 & 20 & 8 & 1 \\ 0 & 2 & 8 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} /(64) \qquad (7)$$

The high pass filter output 304 is applied to an absolute value detector 305 whose output is sent in turn to a statistic estimator 307, which is a shape-adaptive windowing local standard deviation (SD) estimator. The shape adaptive windowing, conceptually based on a homogenous region of similar pixels to the current one in a local window, is required for a reliable local SD estimation in a varying environment in a picture. The shape adaptive windowing segmentation data described further in detail with reference to Figure 4, is composed of, at the input 110, a local binary window, w(i-m, j-n)$\in \{0,1\}$, for the current pixel of coordinates (i, j) and, at the input 111, the number N of "1" for similar pixels to the current pixel in the window. Clearly, N equals to:

$$N = \sum_{(i,j)\in W} w(i-m, j-n) \qquad (8)$$

A standard deviation estimator, such as 307, can be generally based on the following equations:

$$\mu(i,j) = \left(\sum_{(m,n)\in W} w(i-m, j-n) g(i-m, j-n) \right) / N \qquad (9)$$

and,

$$\sigma(i,j) = C\left(\sum_{(m,n)\in W} w(i-m, j-n) \mid g(i-m, j-n) - \mu(i,j) \mid \right) / N \qquad (10)$$

[0054] In Equation (10), g(i, j), μ(i, j) and σ(i, j) are respectively the estimator input, the internal local mean and the estimated local SD output. Moreover, depending on the anticipated noise distribution the constant C can be chosen in accordance with equal to 1.25 appropriately for additive Gaussian noise, to 1.15 for additive uniform noise or, simply omitted. In the preferred embodiment, the window dimension is chosen as 5 lines x 11 columns. For the high frequency signal, the local mean μ(i,j) can be set to zero in Equation (10).

[0055] The SD estimator 307 output, at 308, is provided to a look-up table SD-LUT 309 further controlled by Coding Parameters 102. The purpose of SD-LUT 309 is to convert the estimated local standard deviation at 308 to the standard deviation of an equivalent additive noise. SD-LUT 309 generation is previously described in US Patent Application No. 09/603,364 (now U.S. 6,633,683 issued October 14, 2003) by the present inventors and assigned to the same assignee, which application is incorporated herein by reference. In that application, a generic configuration and method for random and correlated noise reduction are described. SD-LUT 309 estimates at its output 310 a mean value of local noise input SD $\sigma_m(x,y)$ (or variance $\sigma_m^2(x,y)$). The LUT input-output relationship between the two local standard deviations $\sigma_r(x,y)$ (or variance $\sigma_r^2(x,y)$) and $\sigma_m(x,y)$ (or variance $\sigma_m^2(x,y)$) can be described by the following method. Let consider the linear portion of the expression representing weight $K(x,y)$:

$$K(x,y) = \left(\sigma_g^2(x,y) - \sigma_n^2(x,y)\right) / \left(\sigma_g^2(x,y)\right) \qquad (11)$$

wherein the unknown additive noise variance $\sigma_n^2(x,y)$ is expected to be varying. It is thus necessary to pre-estimate this variance value for each pixel located at (x,y).

[0056] Referring now to Figure 7, in many situations where the processing is well defined, such as for NTSC or PAL encoding/decoding and DCT-based compression/decompression, an available original and clean test signal $f(x,y)$ can be used for noise evaluation. Figure 7 illustrates partly a proposed configuration generally designated at 760 used for performing an off-line noise variance pre-estimation. The original test signal $f(x,y)$ at 750 is applied to the above-mentioned processing at 751 that gives a test noisy image signal $g(x,y)$ at 752. The additive test

noise signal $n(x,y)$ at 754 is then obtained by the difference $(g(x,y)-f(x,y))$ provided by an adder 753 and is sent in turn to a statistic calculator 755 similar to the calculator 307 shown in Figure 3. The test noise SD $\sigma_{tn}(x,y)$ (or the test noise variance $\sigma_{tn}^2(x,y)$) estimation is done in the same context as that of the luminance signal Y 100 in LU-LOSEGANR 108 shown in Figure 4, with the segmented window parallel signals w(l-m,j-n) at 110 and the selected-pixels count signal N at 111. That is, for a considered pixel at (x,y), one may obtain a pair of SD values $(\sigma_r(x,y),\sigma_n(x,y))$ (or a pair of variance values $(\sigma_r^2(x,y),\sigma_n^2(x,y))$). For the whole test picture or set of test pictures, a given value of σ_r (or σ_r^2) can have many resulting values of σ_{tn} (or σ_{tn}^2). In order to obtain a unique input-output relationship for the SD-LUT 309, it is necessary, for a given σ_r (or σ_r^2), to define a single value σ_m representing all possible values of σ_m. For the preferred SD calculation, proposed estimations for σ_m are as follows:

$$\sigma_m = \text{mean} (\sigma_{tn}, \text{given a value of } \sigma_r); \qquad (12)$$

or

$$\sigma_m = \text{mode} (\sigma_{tn}, \text{given a value of } \sigma_r) \qquad (13)$$

[0057] The estimation (12) or (13) can be done then on an off-line basis by a data storage and estimation device 757. The input-output result (σ_r,σ_m) at 704 and 758 respectively, permits the establishment of a pre-calculated SD-LUT 309 for real time processing involving an unknown image. If the memory SD-LUT 309 is large enough, some controllable bits can be fed at parameters input 102 representing a learning or functional condition, for example for NTSC, PAL or 12Mbit MPEG. The main requirement of the method is the prior knowledge of the processing to create the noisy image $g(x,y)$ from the clean image $f(x,y)$. In the present case, the SD-LUT 309 is empirically obtained with various test sequences coded by 16 usual bit rates corresponding to end-user controlled Coding Parameters 102. Figure 8, in the preferred embodiment, represents typically the relationship between the observed SD

and the noise coding SD for various Coding Parameters 102. The SD-LUT output 310, designated by $\sigma_m(i,j)$, is applied to the weighting function 311 for MMSE noise reduction explained further with reference to Figure 4. Depending on the detected region at the current pixel location (i,j) indicated by region map 105, the weighting function output signal 312, designated now by $\sigma_e(i,j)$, is empirically given by the following equation:

$$\sigma_e(i,j) = \begin{cases} (4/8).\sigma_q(i,j), & \text{for Edge Region} \\ (5/8).\sigma_q(i,j), & \text{for Near Edge Region} \\ (5/8).\sigma_q(i,j), & \text{for Near Flat Region} \\ (2/8).\sigma_q(i,j), & \text{for Flat Region} \\ (2/8).\sigma_q(i,j), & \text{for Texture Region} \end{cases} \quad (14)$$

It is worthwhile to note that, in the present embodiment, the noise contribution on Edge pixel is considered as important as the noise contribution on Near-Edge or Near-Flat regions. Such noise will be heavily filtered in these three regions. Inversely, the filtering in Texture region should be sufficiently light enough, since texture already masks noise. Finally, in Flat regions, noise is relatively small and nearly random; excessive filtering will degrade eventually fine but visible signal details.

[0058] In order to smooth the region transitions, the weighting function output signal 312 $\sigma_e(i,j)$ is applied to a 2D low pass filter L2 at 313, which is a separable filter. The 2D impulse response is:

$$L_2(i,j) = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} /(16) \quad (15)$$

[0059] The filter output signal $\sigma_n(i,j)$ at 107, considered as local SD of an equivalent additive but varying noise, is provided to the noise correcting block 108.

[0060] Referring now to Figure 4, there is illustrated a block diagram of the LUminance component LOcal SEGmentation-based Adaptive Noise Reducer (LU-LOSEGANR) 108. There are many Spatial Adaptive Noise Reduction techniques known in the art. However, few of them are, firstly, robust in presence of noise and, secondly, efficient in the Edge Region(s) of a picture. LU-LOSEGANR 108 is a simplified version of the generic Adaptive Noise Reducer described in the above-cited

U.S. Patent Application No. 09/603,364. In order to give some robustness to a local segmentation in the presence of noise, a simple low pass filter 401 described by Equation (15) is utilized for the noisy input signal 100. The filter output 402, denoted as g*(i,j) is applied to a local window segmentation 403. The later provides, in the considered window domain W, a set of binary signals w(i-m,j-n) 110 defined as:

$$\text{For } (m,n) \in W,\ w(i-m,j-n) = \begin{cases} 1, & \text{if } |g^*(i-m,j-n) - g^*(i,j)| < \text{Threshold} \\ 0, & \text{otherwise.} \end{cases} \qquad (16)$$

[0061] Thus, the binary signals w designate a homogenous region, within a threshold tolerance, to the current pixel located at (i,j). The local window therefore becomes shape-adaptive. The threshold value is applied at 416 and is set, in the preferred embodiment, to 12. The number N, at 111, of "1" for similar pixels in the window, is provided by the counter 405.

[0062] In order to provide efficient estimation of the two first order signal statistics, the window binary signals 110 and its parameter N signal 111 are connected to a local mean calculator 404 and a local SD calculator 407 for the noisy input signal 100. The calculators are described respectively again by the above equations (9) and (10).

[0063] Finally, in order to provide efficient noise reduction in a varying environment of picture signal, such as edge regions, a MMSE coring technique is given by a gain calculator 411 operating on the two SD values, the first one $\sigma(i,j)$ 408 coming from the noisy signal, the second one $\sigma_n(i,j)$ 107 coming from noise power estimator illustrated in Figure 3. The said MMSE coring K(i,j), at 415, is described by the following equation:

$$K(i,j) = \max\left[0, \frac{\sigma^2(i,j) - \sigma_n{}^2(i,j)}{\sigma^2(i,j)}\right] \qquad (17)$$

A possible simplified version of Equation (17), at the expense of heavier signal reduction, is an MMSE-like coring defined as:

$$K(i,j) = \max\left[0, \frac{\sigma(i,j) - \sigma_n(i,j)}{\sigma(i,j)}\right]. \quad (18)$$

[0064] Finally, the filtered output luminance signal Y*(i,j) at 109 is given by

$$Y^*(i,j) = \text{mean}[Y(i,j)] + K(i,j).\{[Y(i,j) - \text{mean}[Y(i,j)]]\} \quad (19)$$

using a first adder 409 having its output 410 feeding a multiplier 414 receiving K(i,j) at 415 and feeding in turn a second adder 412, as illustrated in Figure 4.

[0065] Referring now to Figure 5 that represents the CHrominance component LOcal Noise power Estimator (CH-LONE) block diagram 112 of Figure 1. The CH-LONE principle is similar to the luminance case. For each chrominance component, as illustrated, CH-LONE 112 comprises a diamond shape low-pass filter 501 for high frequency component extraction with an output 502 connected to a subtractor 503. Subtractor 503 subtracts output 502 from Noisy Cu/Cv 101uv for output 504 to an absolute value detector 505. Local shape adaptive noise power estimator with a 2H up-sampler 507 receives output 506 from detector 505. 1D low-pass filter 509 receives output 508 and supplies its output 510 to a shape-adaptive local standard deviation estimator 511. Output 512 of estimator 511 is provided to a 2H down-sampler 513 and thereafter via output 514 to additive noise SD-LUT 515. SD-LUT output 516 is connected to multiplier 518 that applies a weighting factor 517. The proposed configuration is based on some assumptions: firstly, for simplicity purpose, the shape adaptive local windowing can be the same as in the luminance signal; secondly, for the use of the luminance-based window segmentation data, it has been experimentally found that good results can be obtained if the chrominance is interpolated to the luminance resolution via up-sampler 507 and low-pass filter 509 followed by down-sampler 513. (Conversely, decimating the luminance-based window segmentation data to the chrominance resolution does not yield a better solution); thirdly, it is not necessary to classify the chrominance image to multiple regions as in the luminance case; and finally, in the proposed apparatus and method and as found through experimentation, the weighting factor applied after the SD-LUT 515 by multiplier 518 is sufficiently set to equal to (½) at 517 in order to re-use the same luminance SD-LUT (309).

[0066] Referring now to Figure 6, there is represented a block diagram of the CHrominance component LOcal SEGmentation-based Adaptive Noise Reducer (CH-LOSEGANR) 115. Again, the noise reduction technique in each chrominance component is similar to the technique for luminance noise reduction illustrated in Figure 4.

[0067] The main difference is the appropriate signal used for interpolation by up-sampler 601 and Interpolation filter 603 as required, firstly, for the estimation of the first two statistics using the luminance-based window segmentation data; and secondly, for the Flat region classification as described before with reference to Figure 2. For a 4:2:2 video-sampling pattern, the illustrated by-two (2) up-sampler 601 is simply horizontal, the corresponding interpolator being a horizontal half-band filter. In the proposed system, the filter impulse response is given by the following coefficients: (-5, 0, 37, 64, 37, 0, -5)/(64). Of course, appropriate down- samplers 609 and 610 following, respectively, the local mean calculator 605 and the local standard deviation calculator 606 are necessary for respecting the original chrominance resolution. Since the local mean and the local standard deviation are slowly varying, no filter is required further for these down sampling operations. For a 4:2:0 or other sampling patterns, the up-sampler 601 and the interpolation filtering are more elaborate but well known to people of ordinary skill in the art.

[0068] Moreover, for chrominance video components, even theoretically zero-mean signals, a local mean calculator 605 is still utilized. Its presence can be justified since a local windowed signal mean is not necessary equal to zero. It is interesting to note again that, for noise correction, the luminance-based shape adaptive windowing, previously described, is generally sufficient for chrominance signals.

[0069] While the invention is described with reference to MNR apparatus 50, persons skilled in the art will readily understand that the methods described herein may be embodied in a computer readable medium containing executable instructions for enabling a programmable processor (e.g. complex programmable logic device (CPLD), filed programmable gate array (FPGA), micro processor, etc.) to perform the methods of the invention. Further, the invention herein may comprise a computer system including a processor programmed by such executable instructions.